The Black–Scholes pricing formula in the quantum context

(Brownian motion/quantum process/stochastic integration/Ito lemma/Black–Scholes–Merton theory)

Wiliam Segal*[†] and I. E. Segal‡§

*Office of Policy Development and Research, Department of Housing and Urban Development, 451 7th Street, SW, Room 8212, Washington, DC 20410; and
‡Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 2-244, Cambridge, MA 02139

ABSTRACT A natural explanation for extreme irregularities in the evolution of prices in financial markets is provided by quantum effects. The lack of simultaneous observability of relevant variables and the interference of attempted observation with the values of these variables represent such effects. These characteristics have been noted by traders and economists and appear intrinsic to market dynamics. This explanation is explored here in terms of a corresponding generalization of the Wiener process and its role in the Black–Scholes–Merton theory. The differentiability of the Wiener process as a sesquilinear form on a dense domain in the Hilbert space of square-integrable functions over Wiener space is shown and is extended to the quantum context. This provides a basis for a corresponding generalization of the Ito theory of stochastic integration. An extension of the Black–Scholes option pricing formula to the quantum context is deduced.

In the past several decades, the role of information in economics has expanded greatly, and it now is recognized as a dominant factor in several areas, most notably financial markets. Imperfect information appears especially as the culprit that may defeat earlier rational process analysis. But beyond this, it is clear that there are many contexts in economics in which information is not merely lacking but in which this is inherently and fundamentally the case. The information on several relevant factors—consumer preferences and technological capability, e.g.—may be not simply missing but lacking in simultaneous observability. Attempts to observe the missing information may react on and alter complementary information.

These features characterize a quantum system in the mathematical sense, to which is applicable von Neumann's general treatment (1) of observables and states, as well as the philosophical ideas of "complementarity" due to Bohr and of "indeterminacy" due to Heisenberg. In particular, one of the most evident features of a financial market is the impossibility of observing prices and their instantaneous forward time derivatives. This quite fundamental aspect is not captured in the conventional formalism in which both quantities are represented as random variables. Conventional random variables are in principle simultaneously observable, whereas a lack of simultaneous observability appears capable of precise mathematical formulation only in quantum terms.

The literature treating the general importance of factors outside the realm of public information for market dynamics includes works of traders and economists, e.g., the books of Keynes (2) and Soros (3) and of Grossman (4) and Shiller (5). A quite small sampling of the extensive journal literature on specific aspects includes Delong *et al.* (6), Grossman and Stiglitz (7), Lo and MacKinlay (8), Lo and Wang (9), and Poterba and Summers (10).

The classical (= unquantized) theory of prices in financial markets that originated with Black and Scholes (11) and Merton (12) (hereafter, BSM theory) has been highly successful and is widely accepted. A quantum variant of this theory serves to facilitate the modeling of phenomena not fully explained by it, such as short-term volatility, extreme discontinuities, serial correlation, and notably large-tailed distributions.[¶] At the same time, it provides a formal mathematical basis for the reconciliation in principle of the market rationality hypothesis with empirically based literature such as that cited, and trading practices productive of asymmetric information, such as front-running in one form or another.

The simplest version of such a theory probably consists of the substitution of a quantum process for the Wiener process that underlies the BSM theory. However, this substitution requires the associated development of a theory of integration that appropriately extends the Ito theory (13). This article presents this development and applies it to the derivation of a quantum version of the Black–Scholes option pricing formula.

Random Variables in the Quantum Context

The mathematical foundations of probability theory due to Kolmogoroff represent random variables as measurable functions on a measure space M. The operation of multiplication by a measurable function, acting on the Hilbert space **H** of all square-integrable functions over the measure space, is then a self-adjoint operator in **H**. Conversely, any self-adjoint operator in **H** is (unitarily) equivalent to the operation of multiplication by a measurable function.

The operations of multiplications on **H** by measurable functions on M are mutually commutative. In the quantum context, (generalized) random variables are represented by self-adjoint operators and differ from those of conventional probability theory essentially only in that they need not commute. Indeed, they are mutually commutative if and only if they are simultaneously observable or, in mathematical terms, simultaneously representable as multiplications by measurable functions.

A classic physics example is that of the ensembles of the particle and field strength observables of light. Each ensemble by itself consists of mutually commutative observables, but the two ensembles do not commute with each other. The quantized light field as a unified entity is mathematically perfectly well defined and entirely explicit, notwithstanding its anthropomorphically unintuitive features, and serves to reconcile

---

[†]e-mail: William_Segal@hud.gov.
§To whom reprint requests should be addressed. e-mail: ies@math.mit.edu.

¶For example, a sum of quantum random variables (see below), each of which has the same probability distribution as a classical random variable, and which equally have pairwise vanishing covariances, will generically have more extreme deviations (e.g., have higher even moments beyond the second) than the sum of the corresponding classical variables.

Economic Sciences: Segal and Segal

Proc. Natl. Acad. Sci. USA 95 (1998)   4073

simply and effectively the corpuscular and wave aspects of light.

Probability theory is equivalent in all observable aspects to just that special case of a quantum system, modeled as an ensemble of self-adjoint operators together with an expectation value form on the ensemble, in which there is full simultaneous observability. This follows from commutative spectral theory, according to which any ensemble of commuting self-adjoint operators is unitarily equivalent to a corresponding ensemble of multiplications by measurable functions acting on the space of square-integrable functions over a measure space. An "expectation value functional" E on the ensemble satisfying the basic constraints of linearity and positivity (hereafter, simply "state," for brevity and positivism) then corresponds to integration over the measure space, and conversely.

This formulation of a system of observables may be rationalized by postulates applicable to quantitative science in general and not only to physics. For brevity, however, we define here simply a "quantum system" (A, E, H) as one consisting of a linear ensemble A of self-adjoint operators in the Hilbert space H that is closed under addition and the "Jordan" product $AoB = (AB + BA)/2$, together with a linear functional E on A satisfying the constraints (positivity and normalization) that $E(A^2) \geq 0$ for all A in A, and $E(I) = 1$, where I denotes the identity operator (i.e., E is a state in the sense given above). The Jordan product is used because the product AB of observables A and B is definable in observable terms (and is self-adjoint) only when they are simultaneously observable, which is equivalent to their commutativity. In particular, A^2 is definable, which implies that AoB is definable because it may be expressed as $[(A + B)^2 - A^2 - B^2]/2$. (As argued by von Neumann, the sum $A + B$ is definable by virtue of the linearity of E, which permits $A + B$ to be defined as that observable such that $E(A + B) = E(A) + E(B)$ in all states of the system.)

Quantum effects may be incorporated directly into the BSM analysis by the addition to the Wiener process $W(t)$ representing the evolution of public information affecting the market of a distinct process $X(t)$ representing the influence of factors not contemporaneously observable with those involved in $W(t)$. $X(t)$ may in itself also be equivalent to a Brownian motion, but the conjunction of $W(t)$ with $X(t)$ cannot be represented as simply a two-dimensional Brownian motion process.

Some degree of specification of the nature of the lack of simultaneous observability is naturally required for mathematical development. The absence of simultaneous observability for two observables A and B is represented mathematically by the nonvanishing of the commutator AB-BA. In the absence of *a priori* information on the commutators of the $W(t)$ with the $X(t')$, it appears reasonable to assume that they are constants dependent on t and t′, to a first approximation.

To set up an explicit simultaneous mathematical representation for both processes that incorporates such features, it is natural to use the formalism of the quantum theory of systems of an infinite number of degrees of freedom, i.e., the quantum field theory. The Wiener process as well as the quantum processes treated below may in fact be given simple mathematical forms in these terms. The (free boson) quantum field over a given Hilbert space H (e.g., see ref. 14), consists of a structure (K, Φ, Γ, v) in which K is a Hilbert space, Φ is a mapping from vectors in H to self-adjoint operators in K, Γ is a continuous mapping from the unitary operators on H into those on K, and v is a unit vector in K, which is uniquely characterized by the following properties:

(*i*) The "Weyl" (or "canonical commutation") relations: for arbitrary x, y in H, $\exp(i\Phi(x))\exp(i\Phi(y)) = \exp(i\Phi(x + y))\exp(i\text{Im}<x, y>/2)$.

(*ii*) The liftup of transformations on H to those on K: if U is a unitary operator on H and x is a vector in H, then $\Gamma(U)\Phi(x)\Gamma(U)^* = \Phi(Ux)$; $\Gamma(U)v = v$; and the self-adjoint generator of the unitary one-parameter group $\Gamma(e^{it})$ is positive.

(*iii*) Minimality of K: the linear combinations of the $\exp(i\Phi(x))v$ (x arbitrary in H) are dense in K.

The vector v represents the "vacuum" state of the quantum field, in which the expectation value E(A) of a self-adjoint operator A in K takes the form $<Av, v>$. The vacuum state is invariant under $\Gamma(U)$ for an arbitrary unitary operator U on H, but there are other states with this property of "universal" invariance, including in practice temporal invariance, which is characteristic of an equilibrium state. The most general ergodic (or "pure") universally invariant state (15), denoted E_γ, is that whose generating function $E[\exp(i\Phi(x))]$ takes the form $\exp(-\gamma^2\|x\|^2/4)$, where γ is a constant ≥ 1. ($\gamma = 1$ corresponds to the vacuum state; values of $\gamma < 1$ are not consistent with positivity.)

Temporal invariance considerations are simplified by treating processes on the entire real line, rather than on an interval or half-axis (see, e.g., refs. 16 and 17). A quantum point (resp. interval) process will be defined as a map from a real interval (resp., an additive map from the subintervals thereof) to the operators of a quantum system (A, E, H). For example, if $t \to f_t$ is a map from $\Re$ to functions on $H = L_2(\Re)$, then the map $t \to \Phi(f_t)$ from $\Re$ to self-adjoint operators in the quantum field Hilbert space K is a quantum point process, relative to any given state E. The map $(a, b) \to \Phi(f_b - f_a)$ from intervals (a, b) (with a < b) on $\Re$ to the self-adjoint operators is the associated quantum interval process.

In the state E_γ, if $x_1, \ldots x_n$ is any set of real vectors in H, then the (commutative observables) $\Phi(x_1)$, $\Phi(x_n)$ are normally distributed with vanishing means and covariance matrix $\frac{1}{2}\gamma^2\|<x_i, x_j>\|$). In particular, if c_t denotes the characteristic function of the interval [0, t], then for $t \geq 0$, $\Phi(c_t)$ is equivalent to a Wiener process on the positive half-axis of variance parameter $\frac{1}{2}\gamma^2$. The corresponding interval process is then defined on all finite intervals in $\Re$ and is temporally stationary, as a consequence of the universal invariance of the E_γ. Thus, the unitary transformation U(t): $f(x) \to f(x + t)$ in H induces the unitary transformation $\Gamma(U(t))$ in K, and this is equivalent to the transformation of random variables on Wiener space $x(J) \to x(J + t)$ where $x(J)$ denotes the increment in the Wiener process over the interval J. The space of all square-integrable random variables on Wiener space for the full line is correspondingly equivalent to K by a unitary transformation that carries the function identically 1 on Wiener space into v.

The operators $\Phi(f)$ are mutually commutative if the f are restricted to be real functions in H, and the same is true of the operators $\Phi(if)$, but $[\Phi(f), \Phi(ig)] = i<f, g>$ if f and g are real. Moreover, the $\Phi(ic_t)$ also form a Wiener process apart from nomenclature. Thus, $\Phi(ic_t)$ is a natural candidate for the quantum information process $X(c_t)$ considered above, and the combined processes "$aW(t) + bX(t)$" may be represented as $\Phi((a + ib)c_t)$. The present method is simply applicable to a considerably more general class of processes, including those with serial correlation, such as the Ornstein–Uhlenbeck (O–U) process suggested as a model for certain markets (see, e.g., ref. 9). The next section defines this class and shows that these processes are differentiable in a generalized sense, on which their integration theory is later based.

Differentiation of Quantum Processes

We define a pseudo-Wiener process (hereafter, PSW) as a point map $t \to \Phi(f_t)$, or corresponding interval map $[a, b] \to \Phi(f_b - f_a)$, provided the following conditions are satisfied:

(*i*) The map $t \to f_t$ is continuous from the reals to H;

(*ii*) $<f_t, f_s>$ is left-differentiable as a function of t, for arbitrary fixed s, with a uniformly bounded left derivative;

(*iii*) For all t, $\lim_{h\to 0} h^{-1}\|f_{t+h} - f_t\|^2$ ($\equiv k(t)$) exists;

(*iv*) $<f_t, g>$ is left-differentiable as a function of t for arbitrary functions g in the domain $\mathbf{D_o}$ of infinitely differentiable functions on $\mathfrak{R}$ for which all derivatives are square-integrable.

These conditions are weak in practice and are evidently satisfied in simple cases such as the classical Wiener or O–U processes. In the latter cases, the functions f_t are real-valued whereas in the quantum case they in general are complex-valued. In the application to markets, the real part of f_t may be interpreted as a representation of the public information process, and the imaginary part represents information that is not observable contemporaneously with the public information.

Differentiation and integration theory for quantum processes is facilitated by correlating the process values with sesquilinear forms on a dense domain **D** of sufficiently regular vectors in **K**. Any operator A having **D** in its domain determines (and may thereby be identified with) a form f_A by the equation $f_A(x, y) = <Ax, y>$. In general, sesquilinear forms on **D** will not arise from operators in this way, and their product is undefined. However, forms have adjoints, and their products with an operator (in either order) is defined, provided the operator and its adjoint leave **D** invariant. Thus, $f*(x, y) = f(y, x)*$, where for a complex number a, a* denotes the complex conjugate; and $(fB)(x, y) = f(Bx, y)$, $(Bf)(x, y) = f(x, B*y)$, if f is a given form and B a given operator. Operators may thereby be naturally identified with a special class of forms.

$\Phi(f)$ is an unbounded operator except when $f = 0$, and its domain requires constraint. Let $\mathbf{D_1}$ denote the domain in **K** consisting of the linear combinations of the vectors of the form $\Phi(f_1) \ldots \Phi(f_n)v$, where the f_j are in $\mathbf{D_o}$. Any PSW V(t) determines a corresponding form-valued function ϕ_t of t by the equation $\phi_t(x, y) = <V(t)x, y>$, and this function is differentiable. This implies in particular the differentiability of the Wiener process as a form on $\mathbf{D_1}$. However, $\mathbf{D_1}$ need not be invariant under the V(t) or smooth functions thereof. To attain this property, $\mathbf{D_1}$ will be enlarged to the domain **D** consisting of all linear combinations of the $F_1(V(t_1)) \ldots F_n(V(t_n))y$ where y is arbitrary in $\mathbf{D_1}$ and the F_j are in the space **F** of C^3 functions on $\mathfrak{R}$, which, together with their first three derivatives, are $O(\exp(x^a))$ with $a < 2$. In **D**, differentiability is only one-sided, but this is sufficient for present purposes.

THEOREM 1. $<V(t)u, u'>$ *is a left-differentiable function of* t, *for arbitrary fixed* u *and* u' *in* **D**.

The proof uses the fact that $E(AB) = E(A)E(B)$ if A and B are functions of the $\Phi(x)$ with the x ranging over respective orthogonal submanifolds of **H**. It uses also the simultaneous spectral resolution of the F(x) for real x, in terms of an infinite sequence of independent identically distributed Gaussian random variables (14). The complex vectors in $\mathbf{D_o}$ may be eliminated from the definition of $\mathbf{D_1}$ by repeated use of the canonical commutation relations and the fact that, for real f, $\Phi(if)v = -i\Phi(f)v$. Given u and u' in D_1 may be expressed in terms of the application of Φ to a finite-dimensional submanifold of **H**. Applying the Gram–Schmidt process to a basis for the manifold spanned by this submanifold together with the f_t then expresses $<V(t)u, u'>$ as a linear combination of Gaussian integrals whose only dependence on t is via external coefficients of the form $<f_t, g>$ with g either in $\mathbf{D_o}$ or of the form f_s.

The boundedness and the continuity of the left derivatives of the $<f_t, g>$, except at the point $t = s$ when $g = f_s$, implies also

COROLLARY 1. *For arbitrary* u *and* u' *in* **D**, $<W(t)u, u'>$ *is a continuous as well as left-differentiable function of* t. *Moreover, the left derivative is uniformly bounded and continuous with the possible exception of a finite number of* t-*values in any finite interval.*

Integration of Quantum Processes

For a given PSW W(t), the left derivative of W(t) as a form on the domain **D** will be denoted as w(t). The product of the form w(t) with the self-adjoint operator F(W(t)) leaving **D** invariant is defined for F in **F**, making possible the definition of the integral $\int_0^t F(W(s))w(s)ds$ as the weak integral J given by the equation $<Ju, u'> = \int<F(W(s))w(s)u, u'>ds$.

The quantum extension of Ito's formula for F(W(t)) involves the derivative of F as a function of noncommuting operators, which is in general not simply F'. To avoid technicalities extraneous to ideational aspects of financial markets, F will now be limited to be of the form $F(x) = \int\exp(ax)dG(a)$, where the integration is over all of $\mathfrak{R}$, and G is locally of bounded variation and such that the integrals representing F, F', and F'' are absolutely convergent. This class of functions will be denoted as $\mathbf{F_o}$, and the domain **D** will be redefined by restricting the functions F in the earlier definition of **D** to be in $\mathbf{F_o}$. The following result extends Ito's lemma for this extensive and adequate class of functions and implies that the indicated weak integral effectively coincides with stochastic integration in the usual sense.

THEOREM 2. *Let* F *be arbitrary in* $\mathbf{F_o}$. *Then* $F(W(t)) = F(0) + \int_0^t [F'(W(s))ow(s) + k(s)F''(W(s))]ds$.

The proof follows from the theorem of Scheefer (18) according to which continuous functions whose left derivatives coincide with the possible exception of a countable set of points differ only by a constant. Each side of the foregoing equation is both continuous and left differentiable as a function of t. To evaluate the derivative of the left side, consider first the case $F(x) = \exp(ax)$. The corresponding difference quotient simplifies via the extension of the Weyl relations to the $\exp(\Phi(x))$, which are unbounded but defined on **D**. Wick's theorem on the expectation values of products of quantum fields extends to the states E_γ and shows that effects of order higher than the second give vanishing contributions. The derivative of the right side is immediate from the definition of the integral and coincides with that of the left side. The case of general F in $\mathbf{F_o}$ follows from the linearity in F of both sides of the conclusion.

Example: The conventional equation for the temporal evolution of a stock price P(t) requires only the use of the Jordan product to adapt to the present context, together with the use of forms. The stochastic exponential of a quantum process P(t) may be defined as the solution of the integral equation $E(t) = 1 + \int_0^t P'(s)oE(s)ds$. Theorem 2 shows that, if $P(t) = r + w(t)$, then $E(t) = \exp[rt + W(t) - \frac{1}{2}kt]$.

The Black–Scholes Formula

The derivation in the present context is based on the usual simplifying assumptions regarding market dynamics, notably the absence of riskless arbitrage possibilities. It is assumed also that there is no friction or dividends, that there is a riskless interest rate r, and that short selling is unrestricted, trading is continuous, and securities are infinitely divisible. For simplicity, the basic parameters are assumed constant (the method extends to the case when they are given as functions of t). Stock prices are assumed to satisfy the equation of Corollary 2.

The Feynman–Kac formula in the noncommutative form based on ref. 19 is then applicable by the same argument as in the classical case (20). If g is the terminal value function, r is the riskless interest rate, T the elapsed time, and P(t) the stock price process, the value V of the option takes the natural form of the expected discounted payoff, after replacement of the stock price process with a conceptual process having the modified drift rP(t): $V = E_d[e^{-rT}g(P(T))]$.

The normal distribution of $\Phi(x)$ whose characteristic functions is given above implies a corresponding lognormal distribution for the price change over the interval [0, T] of mean

Economic Sciences: Segal and Segal

Proc. Natl. Acad. Sci. USA 95 (1998)   4075

$T(r - k/2)$ and variance $\frac{1}{2}g^2\|f_T\|^2$, noting that $\Phi(x)$ is normally distributed with vanishing mean and variance $\frac{1}{2}g^2\|x\|^2$. V may accordingly be expressed as a Gaussian integral, which in the case of a European call option with initial price P and strike price K takes the following form, in terms of the parameter $s = \|f_T\|/T^{1/2}$ generalizing the classical volatility:

$$V = P \exp[(1/2)T(\gamma^2 s^2 - k)]N[\{\log(P/K) + T(r + \gamma^2 s^2$$
$$- k/2)\}/\gamma s T^{1/2}] - Ke^{-rT}N[\{\log(P/K)$$
$$+ T(r - k/2)\}/\gamma s T^{1/2}].$$

where $N(x) = (2\pi)^{-1/2}\int_0^x \exp(-y^2/2)dy$.

The classical formula is obtained by the substitutions $\gamma = 1$, $s = \sigma$, $k = \sigma^2$. The formula involves two additional parameters beyond the volatility, in the usual case in which $f_t = c_t$, but γ and s occur only in the combination γs. They could be estimated separately by a determination of the expectation functional for prices conditional on an objective quantitative designation of "public" information. For models with serial correlation, k typically vanishes. It should be interesting to test empirically whether in fact $k = s^2$ as in the Wiener process model or $k = 0$ as typically in processes with serial correlation.

The results for options involving multiple times or time-dependent parameters may be expected to show greater differences between the classical and quantum cases because of the generic noncommutativity of the quantum prices at several future times. Analytic estimates show that when T is sufficiently large, increases in γ while holding the other parameters constant correspond to a greater prevalence of irregularities.

Summary

The quantum extension of the Black–Scholes–Merton theory provides a rational, scientifically economical, and testable model toward the explanation of market phenomena that show greater extreme deviations than would be expected in classical theory. The volatility estimated from daily prices on the time of months may be expected to correspond roughly to γs. γ is interpretable as a measure of market entropy, deriving from factors separate from price volatility, e.g., trading volume. On the longer term, larger values of γ correspond to increased dispersion, as expected.

1. von Neumann, J. (1936) *Sbornik N.s.* **1,** 415–484.
2. Keynes, J. M. (1936) *The General Theory of Employment, Interest, and Money* (Harcourt Brace, New York).
3. Soros, G. (1987) *The Alchemy of Finance* (Simon & Schuster, New York).
4. Grossman, S. J. (1989) *The Informational Role of Prices* (MIT Press, Cambridge, MA).
5. Shiller, R. J. (1989) *Market Volatility* (MIT Press, Cambridge, MA).
6. DeLong, J., Shleifer, A., Summers, L. H. & Walddmann, R. J. (1989) *J. Finance* **44,** 681–696.
7. Grossman, S. J. & Stiglitz, J. E. (1976) *Am. Econ. Rev.* **77,** 246–253.
8. Lo, A. W. & MacKinlay, A. C. (1988) *Rev. Fin. Stud.* **1,** 41–66.
9. Lo, A. W. & Wang, J. (1995) *J. Finance* **50,** 87–129.
10. Poterba, J. M. & Summers, L. H. (1988) *J. Finan. Econ.* **22,** 27–60.
11. Black, F. & Scholes, M. (1973) *J. Political Econ.* **3,** 637–654.
12. Merton, R. C. (1973) *Bell J. Econ. Man. Sci.* **4,** 141–183.
13. Ito, K. (1951) *Mem. Am. Math. Soc.* **4,** 1–51.
14. Segal, I. E. (1956) *Trans. Am. Math. Soc.* **18,** 160–175.
15. Segal, I. E. (1962) *Illinois J. Math.* **6,** 500–523.
16. Kakutani, S. (1950) *Proc. Natl. Acad. Sci. USA* **36,** 319–323.
17. Doob, J. L. (1965) *Stochastic Processes* (Wiley, New York).
18. Hobson, E. W. (1927) *The Theory of Functions of a Real Variable* (Cambridge Univ. Press, Cambridge, MA), Vol. 1.
19. Arveson, W. (1983) CBMS Regional Conference No. 55 (Am. Math. Soc., Providence, RI), pp. 29–40.
20. Duffie, D. (1987) *J. Econ. Theory* **46,** 194–204.